

# The Distribution of Eigenvalues of Covariance Matrices of Residuals in Analysis of Variance

John Mandel

Institute for Materials Research, National Bureau of Standards, Washington, D.C. 20234

(April 9, 1970)

A rigorous definition is given for the concept of an "interaction matrix"  $(Z_{ij})$  where  $i = 1$  to  $m$  and  $j = 1$  to  $n$ , in terms of two idempotent matrices  $A_r$  and  $B_s$  of rank  $r$  and  $s$ , respectively. It is then shown that the frequency distribution of the eigenvalues of  $(Z)(Z)'$  depends only on  $r$  and  $s$ . Applications are given to matrices of residuals arising from two-way data, either by removing row and/or column-means, or by applying any number of sweeps of the "vacuum cleaner." The theorems are important in the theory of the analysis of two-way tables of nonadditive data.

Key words: Analysis of variance; covariance matrix; eigenvalues; interaction; matrix; residuals; two-way table; vacuum cleaner.

## 1. Introduction

In a recent paper [1],<sup>1</sup> a method has been presented for a partitioning of the row by column interaction in two-way tables and the practical usefulness of this method as a tool in data analysis has been demonstrated. The method required the calculation, by Monte Carlo techniques, of quantities that are analogous to the "degrees of freedom" in ordinary analysis of variance. These new degrees of freedom were calculated as follows.

Let  $(Z_{ij})$  be an  $m \times n$  matrix of independent random normal deviates (zero mean, unit variance). Consider the  $m \times n$  matrix of residuals  $(d_{ij})$ , where

$$d_{ij} = Z_{ij} + Z_{..} - Z_{i.} - Z_{.j},$$

where a dot indicates averaging over the index replaced by the dot.

Then the new quantities, denoted as degrees of freedom, are the expected values of the eigenvalues of the matrix product  $(d)(d)'$  (where ' indicates the transpose of a matrix). The expected values of these eigenvalues are functions of  $m$  and  $n$ , and were tabulated as such. Professor John Tukey, of Princeton University, has pointed out that the degrees of freedom calculated by this procedure are of far more general applicability than indicated in the paper in question. The present paper provides the theoretical development from which the more general results are obtained. It is based on a new, mathematically rigorous definition for an interaction matrix, from which an important general result can be derived.

## 2. A Definition of Interaction of Two-Way Tables

Let  $(d_{ij})$  be an  $m \times n$  matrix of normally distributed variates, all of zero mean; and let the covariance of  $d_{ij}$  and  $d_{i'j'}$  be denoted by  $\text{cov}(d_{ij}, d_{i'j'})$ ,

We will say that  $(d_{ij})$  is in an interaction matrix of  $r$  degrees of freedom by  $s$  degrees of freedom if  $\text{cov}(d_{ij}, d_{i'j'})$  is of the form

$$\text{cov}(d_{ij}, d_{i'j'}) = [a_{ii'} \cdot b_{jj'}] \sigma^2,$$

<sup>1</sup> Figures in brackets indicate the literature references at the end of this paper.

where the  $m \times m$  matrix  $A_r \equiv (a_{ii'})$  is idempotent of rank  $r$ , and the  $n \times n$  matrix  $B_s \equiv (b_{jj'})$  is idempotent of rank  $s$ .

We will say that  $A_r$  and  $B_s$  are the covariance matrices associated with the interaction matrix  $(d_{ij})$ .

### 3. The Eigenvalues of $(d)(d)'$

**THEOREM 1:** If  $(d_{ij})$  is an  $m \times n$  interaction matrix with associated covariance matrices  $A_r$  and  $B_s$ , then the nonzero eigenvalues of the matrix product  $(d)(d)'$  (where  $(d)'$  denotes the transpose of  $(d)$ ) are the same as those of the matrix  $(t)(t)'$ , where  $(t_{ij})$  is an  $r \times s$  matrix of normally and independently distributed variates  $t_{ij}$ , of zero mean and common variance  $\sigma^2$ .

**PROOF:** Let  $P = (p_{ii'})$  represent the  $m \times m$  orthogonal matrix whose rows are the eigenvectors of  $A_r$ . Since  $A_r$  is idempotent and of rank  $r$ , we have

$$PA_rP' = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \quad (2)$$

where  $I_r$  is the identity matrix of rank  $r$ . Similarly, representing by  $Q = (q_{jj'})$  the  $n \times n$  orthogonal matrix whose rows are the eigenvectors of  $B_s$ , we have

$$QB_sQ' = \begin{pmatrix} I_s & 0 \\ 0 & 0 \end{pmatrix}. \quad (3)$$

From (2) and (3) we derive, respectively,

$$\sum_k \sum_l p_{ik} p_{i'l'} a_{kl} = \begin{cases} \delta_{ii'} & \text{for } i, i' \leq r \\ 0 & \text{for } i \text{ or } i' > r \end{cases} \quad (4)$$

and

$$\sum_u \sum_v q_{ju} q_{j'v} b_{uv} = \begin{cases} \delta_{jj'} & \text{for } j, j' \leq s \\ 0 & \text{for } j \text{ or } j' > s, \end{cases} \quad (5)$$

where  $\delta_{kl}$  is the Kronecker delta ( $\delta_{kl} = 1$  for  $k = l$ , and  $\delta_{kl} = 0$  for  $k \neq l$ ).

Consider now the transformation

$$(t_{ij}^*) = P(d_{ij})Q'. \quad (6)$$

Then,

$$(t^*)(t^*)' = P(d)Q'Q(d)'P'$$

or

$$(t^*)(t^*)' = P(d)(d)'P'. \quad (7)$$

From (7) it follows that  $(t^*)(t^*)'$  and  $(d)(d)'$  have the same nonzero eigenvalues. Furthermore, we have, on account of (6)

$$t_{ij}^* = \sum_k^m \sum_u^n p_{ik} q_{ju} d_{ku}. \quad (8)$$

Consequently,

$$\text{cov}(t_{ij}^*, t_{i'j'}^*) = \sum_k^m \sum_u^n \sum_l^m \sum_v^n p_{ik} q_{ju} p_{i'l} q_{j'v} \text{cov}(d_{ku}, d_{lv}). \quad (9)$$

Introducing (1) into (9), we obtain

$$\begin{aligned} \text{cov}(t_{ij}^*, t_{i'j'}^*) &= \sigma^2 \sum_k^m \sum_u^n \sum_l^m \sum_v^n p_{ik} q_{ju} p_{i'l} q_{j'v} a_{kl} b_{uv} \\ &= \sigma^2 \left( \sum_k^m \sum_l^m p_{ik} p_{i'l} a_{kl} \right) \left( \sum_u^n \sum_v^n q_{ju} q_{j'v} b_{uv} \right) \end{aligned}$$

which, as a result of eqs (4) and (5), becomes:

$$\text{cov } (t_{ij}^*, t_{i'j'}^*) = \begin{cases} \sigma^2 \cdot \delta_{ii'} \cdot \delta_{jj'} & \text{for } \begin{cases} i, i' \leq r \\ j, j' \leq s \end{cases} \\ 0 & \text{for } \begin{cases} i \text{ or } i' > r \\ j \text{ or } j' > s \end{cases} \end{cases} \quad (10)$$

From (8) it also follows that the  $t_{ij}^*$  are normally distributed with  $E(t_{ij}^*) = 0$  for all  $i$  and  $j$ . Consequently, because of (10), the  $t_{ij}^*$  are equal to zero for all  $i, i' > r$  or  $j, j' > s$ . For  $i, i' \leq r$  and  $j, j' \leq s$ , it follows from (10) that the  $t_{ij}^*$  are normally and independently distributed with zero mean and variance  $\sigma^2$ .

We have already shown that  $(d)(d)'$  and  $(t^*)(t^*)'$  have the same nonzero eigenvalues. If we omit from the matrix  $(t^*)$ , all elements for which  $i > r$ , or  $j > s$  (all of which are zero), we obtain an  $r \times s$  matrix, say  $(t_{ij})$ , such that  $(t)(t)'$  has the same nonzero eigenvalues as  $(t^*)(t^*)'$ , and consequently as  $(d)(d)'$ . QED.

NOTE: We have assumed, in the definition of an interaction matrix, that the  $d_{ij}$  are normally distributed, and have proved that *under this assumption*, the  $t_{ij}^*$  defined by eq (8) (except for those that are identically zero), are *independently and normally* distributed with zero mean and variance  $\sigma^2$ . We can generalize the definition of an interaction matrix by requiring only that the  $d_{ij}$  be *conditionally* normal, given the  $a_{ii'}$  and  $b_{jj'}$ . Since the distribution of the  $t_{ij}$  does not depend on the  $a_{ii'}$  and  $b_{jj'}$ , it is *unconditionally* normal with zero mean and variance  $\sigma^2$ , and Theorem I remains true under the more general definition of an interaction matrix.

#### 4. Residuals from an Additive Structure

THEOREM II: Let  $(d_{ij})$  be a matrix whose elements are

$$d_{ij} = y_{ij} - \hat{\mu} - \hat{\rho}_i - \hat{\gamma}_j \quad \begin{cases} i = 1 \text{ to } m \\ j = 1 \text{ to } n \end{cases} \quad (11)$$

where  $y_{ij}$  are independent normal variates, such that  $E(y_{ij}) = \mu + \rho_i + \gamma_j$ , and whose common variance is  $\sigma^2$ , and where  $\hat{\mu}$ ,  $\hat{\rho}_i$ ,  $\hat{\gamma}_j$  are the usual estimates of the grand mean, the row effects and the column effects. Then  $(d_{ij})$  is an interaction matrix of  $(m-1)$  degrees of freedom by  $(n-1)$  degrees of freedom.

PROOF: It is readily shown from (11) that

$$\text{cov } (d_{ij}, d_{i'j'}) = \left( \delta_{ii'} - \frac{1}{m} \right) \left( \delta_{jj'} - \frac{1}{n} \right) \sigma^2. \quad (12)$$

The matrix  $A = \left( \delta_{ii'} - \frac{1}{m} \right)$  is idempotent. Its trace is  $\sum_i \left( \delta_{ii} - \frac{1}{m} \right) = m - 1$ . Hence its rank is  $m - 1$ .

Similarly, the matrix  $B = \left( \delta_{jj'} - \frac{1}{n} \right)$  is idempotent, of rank  $n - 1$ . This proves the theorem.

#### 5. Residuals from the Vacuum Cleaner<sup>2</sup>

LEMMA I: Let  $(y_{ij})$  be an  $m \times n$  matrix of observations  $y_{ij}$ , with common variance  $\sigma^2$ . Let  $(d_{ij,k})$  be the matrix of residuals obtained from the  $y_{ij}$  after extraction of the grand mean ( $\hat{\mu}$ ), the row effects ( $\hat{\rho}_i$ ), the column effects ( $\hat{\gamma}_j$ ), and  $k$  sweeps of the vacuum cleaner [3]. Then the following relation holds:

$$d_{ij,k} = y_{ij} - (\hat{\mu} + \hat{\rho}_i + \hat{\gamma}_j) - \sum_{l=1}^k [K_l p_{li} q_{lj} + \alpha_{li} q_{lj} + \beta_{lj} p_{li}] \quad (13)$$

<sup>2</sup> The results contained in Lemmas I, II, and III are not novel. They have been derived previously, by a different method, by W. L. Nicholson [2], and are presented here only for completeness.

in which

$$(1) \left\{ \begin{array}{l} \sum_i p_{li} = \sum_j q_{lj} = 0 \\ \sum_i p_{li}^2 = \sum_j q_{lj}^2 = 1 \end{array} \right\} \quad \text{for all } l.$$

$$(2) \quad p_{li} = \frac{\rho_i}{\sqrt{\sum_i \rho_i^2}} \quad q_{lj} = \frac{\gamma_j}{\sqrt{\sum_j \gamma_j^2}}$$

$$(3) \quad p_{li} = \frac{\alpha_{l-1,i}}{\sqrt{\sum_i \alpha_{l-1,i}^2}}, \quad q_{lj} = \frac{\beta_{l-1,j}}{\sqrt{\sum_j \beta_{l-1,j}^2}}$$

$$(4) \left\{ \begin{array}{l} K_l = \sum_i \sum_j d_{ij,l-1} p_{li} q_{lj} \\ \alpha_{li} = \sum_j d_{ij,l-1} q_{lj} - K_l p_{li} \\ \beta_{lj} = \sum_i d_{ij,l-1} p_{li} - K_l q_{lj} \end{array} \right.$$

These relations result from the definition of the vacuum-cleaner process (see [3]), according to which

$$d_{ij,l} = d_{ij,l-1} - [K_l p_{li} q_{lj} + \alpha_{li} q_{lj} + \beta_{lj} p_{li}], \quad (14)$$

and the quantities  $K, p, q, \alpha, \beta$ , are calculated as above.

LEMMA II: Using the notation of Lemma I, we have

$$\left. \begin{array}{l} \sum_i p_{li} p_{l'i} = \delta_{ll'} \end{array} \right\} \quad (15)$$

and

$$\left. \begin{array}{l} \sum_j q_{lj} q_{l'j} = \delta_{ll'} \end{array} \right\} \quad \text{for all } l, l' \quad (16)$$

PROOF: Let  $(t) = (t_1, t_2, \dots, t_n)$  be any vector of  $n$  elements, such that  $\sum_j t_j = 0$ , and assume that the following relations hold for some given value of  $l$ .

$$(1) \sum_j t_j q_{l+1,j} = 0$$

for all  $i$ ,

$$(2) \sum_j t_j d_{ij,l} = 0$$

where  $d_{ij,l}$  satisfies eq. (14) of Lemma I.

Then the following relations hold:

$$(a) \sum_j t_j q_{l+2,j} = 0$$

$$(b) \sum_j t_j d_{ij,l+1} = 0 \quad \text{for all } i.$$

PROOF:

$$\begin{aligned}
 (a) \sum_j t_j q_{l+2,j} &= \sum_j t_j \frac{\beta_{l+1,j}}{\sqrt{\sum_j \beta_{l+1,j}^2}} = \frac{1}{\sqrt{\sum_j \beta_{l+1,j}^2}} \sum_j t_j \left[ \sum_i d_{ij} p_{l+1,i} - K_{l+1} q_{l+1,j} \right] \\
 &= \frac{1}{\sqrt{\sum_j \beta_{l+1,j}^2}} \left[ \sum_i p_{l+1,i} \left( \sum_j t_j d_{ij} \right) - K_{l+1} \sum_j t_j q_{l+1,j} \right] = 0, \text{ because of (1) and (2).}
 \end{aligned}$$

$$\begin{aligned}
 (b) \sum_j t_j d_{ij,l+1} &= \sum_j t_j d_{ij,l} - K_{l+1} p_{l+1,i} \sum_j t_j q_{l+1,j} \\
 &\quad - \alpha_{l+1,i} \sum_j t_j q_{l+1,j} - p_{l+1,i} \sum_j t_j \beta_{l+1,j} = 0, \text{ because of (1), (2), and (a).}
 \end{aligned}$$

Thus, if conditions (1) and (2) are fulfilled for any value of  $l$ , they are also fulfilled for any value  $l' > l$ . In that case, we have

$$\sum_j t_j q_{l'j} = 0 \quad \text{for } l' \geq l+1.$$

If we now make  $t_j \equiv q_{lj}$ , we can verify that conditions (1) and (2) are fulfilled. Hence,

$$\sum_j q_{lj} q_{l'j} = 0 \quad \text{for } l' \neq l,$$

and by definition  $\sum_j q_{lj} q_{lj} = 1$ . The relation  $\sum_i p_{li} p_{li} = \delta_{ll'}$  is proved in a similar way.

LEMMA III: Defining  $(d_{ij,k})$  as in Lemma I, we have

$$\text{cov}(d_{ij,k}, d_{i'j',k}) = \sigma^2 \left( \delta_{ii'} - \frac{1}{m} - \sum_{l=1}^k p_{li} p_{li'} \right) \left( \delta_{jj'} - \frac{1}{n} - \sum_{l=1}^k q_{lj} q_{lj'} \right).$$

PROOF: The lemma is easily verified for  $k=1$ . It is also readily shown, using Lemma II, that if the lemma holds for  $k$ , it also holds for  $k+1$ . Thus, the lemma is proved by mathematical induction on  $k$ .

THEOREM III: Define  $(d_{ij,k})$  as in Lemma I, and assume furthermore that the  $y_{ij}$  are normally and independently distributed, with  $E(y_{ij}) = \mu + \rho_i + \gamma_j$ , and common variance  $\sigma^2$ ; then the  $m \times n$  matrix  $(d_{ij,k})$  is an interaction matrix of  $(m-k-1)$  by  $(n-k-1)$  degree of freedom.

PROOF: The theorem results at once from Lemma III and the fact that a matrix of the form

$$\left( \delta_{ii'} - \frac{1}{m} - \sum_{l=1}^k p_{li} p_{li'} \right),$$

with

$$\sum_i p_{li} p_{li} = \delta_{ll'},$$

is idempotent and of rank  $m-k-1$ . (The trace of this matrix is

$$\sum_i \left[ \delta_{ii} - \frac{1}{m} - \sum_{l=1}^k p_{li}^2 \right] = m - m \left( \frac{1}{m} \right) - k = m - k - 1.)^3$$

<sup>3</sup> The distribution of the  $d_{ij,k}$  is only conditionally normal, given all  $p_{li}$  and  $q_{lj}$  values, but as shown on p. 151, this does not invalidate Theorem I. Consequently, Theorem III is true.



**THEOREM IV:** Let  $(y_{ij})$  be an  $m \times n$  matrix of normal, independent variates  $y_{ij}$ , whose means are  $E(y_{ij}) = \mu = \rho_i + \gamma_j$  and whose common variance is  $\sigma^2$ . Extract from  $y_{ij}$  the mean  $\hat{\mu}$ , the row effect  $\hat{\rho}_i$ , and the column effect  $\hat{\gamma}_j$ , as well as  $k$  sweeps of the vacuum cleaner, and denote by  $(d_{ij,k})$  the matrix of residuals resulting from this treatment.

Then the nonzero eigenvalues of the matrix product  $(d_{ij,k})(d_{ij,k})'$  are the same as those of  $(t)(t)'$  where  $(t)$  is an  $(m-k-1)$  by  $(n-k-1)$  matrix  $(t_{ij})$  of normal, independent variates  $t_{ij}$  of zero mean and common variance  $\sigma^2$ .

The theorem holds for all values of  $k$ , from  $k=0$  to  $k = \min [(m-1), (n-1)]$ .

**PROOF:** The theorem is an immediate consequence of Theorems I and III.

## 6. Principal Result and Summary

As indicated in the introduction, the present work was motivated by a desire to generalize results obtained in connection with a new method of data analysis. We are now in a position to state this generalization in the following precise form.

**THEOREM V:** Let  $(d_{ij})$  be an  $m \times n$  interaction matrix of  $r$  degrees of freedom by  $s$  degrees of freedom. Then the probability distribution of the eigenvalues of the matrix product  $(d)(d)'$  depends only on  $r$  and  $s$ , and not on  $m$  and  $n$ .

**PROOF:** The theorem follows at once from Theorem I.

Our principal result is the following: To study the probability distribution of the eigenvalues of  $(d)(d)'$ , where  $(d)$  is *any* interaction matrix with  $r$  by  $s$  degrees of freedom, it is sufficient to study the probability distribution of the eigenvalues of  $(t)(t)'$ , where  $(t)$  is an  $r$  by  $s$  matrix whose elements are independent random normal deviates (mean zero, unit variance). The probability distribution in question is completely determined by the values of  $r$  and  $s$ .

---

The author is greatly indebted to Professor John Tukey for pointing out to him the generality of his previous results and thereby providing the incentive for the present work.

## 7. References

- [1] Mandel, John, The partitioning of interaction in analysis of variance, Nat. Bur. Stand. (U.S.), 73B (Math. Sci.) No. 4, 309-328 (Oct.-Dec. 1969).
- [2] Nicholson, W. L., Notes for a Series of Seminars on Tukey's Vacuum Cleaner (unpublished), Battelle Memorial Institute, Pacific Northwest Laboratory (April 1, 1967).
- [3] Tukey, J. W., The future of data analysis, Annals of Mathematical Statistics **33**, No. 1, 1-67 (1962).

(Paper 74B3-326)